

University of Groningen

Development of an integrated system for activity-based profiling of matrix metallo-proteases

Freije, Jan Robert

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2006

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Freije, J. R. (2006). *Development of an integrated system for activity-based profiling of matrix metallo-proteases*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

CHAPTER 1

Introduction

An adapted version of Chapter 1.2 has been published: Freije, J.R.; Bischoff, R. *Drug Discov. Today Tech.* **2006**, 3, (1), 5-11

1.1 Matrix metallo-proteases

1.1.1 Classification of the matrix metallo-protease family

Proteases can be classified into four major classes: serine-, cysteine-, aspartic- and metallo-proteases based on their residue or cofactor crucial in catalysis. Most of the zinc-containing metallo-proteases exhibit a characteristic HEXXH (single letter amino acid code, with X being any residue) consensus sequence. The two histidine (H) residues serve as zinc ligands and the glutamic acid (E) residue polarizes a water molecule involved in nucleophilic attack at the scissile peptide bond of the substrate. Matrix metallo-proteases (MMPs) are a subfamily of the metzincins which share the extended zinc-binding HEXXHXXGXXH consensus motif and a conserved methionine 1,4-turn adjacent to the catalytic zinc ion, and also include the adamalysins, astacins and serralysins. The third histidine acts as a third zinc ligand instead of a more distant glutamic acid in most other metallo-proteases [1].

The first animal MMP activity was discovered in 1962 by the experiments of

Gross and Lapière, who studied the biochemistry of tadpole tail metamorphosis and observed that native radioactively labelled collagen fibrils could be dissolved in a tissue culture of tadpole skin [2]. The 23 human MMPs which have been identified at the present time are structurally and functionally related zinc- and calcium dependent enzymes. Based on their domain organization and substrate specificity, MMPs are divided into subgroups of collagenases, stromelysins, matrilysins, gelatinases, membrane-type MMPs (MT-MMPs) and other MMPs. A typical MMP consists of a propeptide of about 80 amino acids, a catalytic metalloprotease domain of about 170 amino acids, a linker peptide of variable length (also called hinge region) and a hemopexin (Hpx) domain of about 200 amino acids. The Hpx domain is a substrate specificity determinant, and in the case of MMP-2, -9 and -13, mediates the interaction with their endogenous inhibitors, the tissue inhibitors of metalloproteases (TIMPs). The Hpx domain is absent in MMP-7, -23 and -26. MMP-2 and -9 (gelatinase A and B), in addition have three repeats of a fibronectine type II motif in the catalytic domain. The MT-MMPs include four type I transmembrane proteins and two glycosylphosphatidylinositol-anchored proteins. The exact domain organization of all MMPs has been extensively reviewed elsewhere and will not be described in detail here [3]. Besides the zinc binding motif in the catalytic domain, the "cysteine switch" motif PRCGXPD in the propeptide is a common structural MMP feature, where the cysteine in the propeptide coordinates the catalytic zinc ion together with the histidine residues of the zinc binding motif. This Cys-Zinc coordination keeps proMMPs inactive by preventing a water molecule, essential for catalysis, from binding to the zinc atom. The catalytic domain has an additional structural zinc ion and 2-3 calcium ions which are required for stability and activity.

Expression of MMP genes can be induced by effectors such as growth factors, hormones, cytokines, physical stress and oncogenic cellular transformation, but also cell-matrix and cell-cell interactions have been identified as inducers of MMP gene expression [4]. Reactive oxygen and nitrogen species have also been shown to be able to induce MMP production by affecting MMP gene expression [11].

MMPs are extracellular proteins, but recent studies have indicated that MMP-1, MMP-2 and MMP-11 are also found intracellular and may act on intracellular proteins [5–7]. MMPs are synthesized as pre-proenzymes. The signal peptide is removed during translation and proMMPs are generated. Most MMPs are secreted as inactive proenzymes, which can be activated upon proteolytic removal of the propeptide domain causing a disruption of the Cys-Zinc interaction [8]. The Hpx domain has been shown to play an important role in activation of proMMP-2 on the cell surface, by binding to the non-inhibitory domain of TIMP-2. This

proMMP-2 - TIMP-2 complex in turn binds to an active MT1-MMP through interaction with the free MMP inhibitory domain of TIMP-2. A second active MT1-MMP subsequently binds to the first MT1-MMP through their Hpx domains, thus forming a tetrameric complex, in which the second MT1-MMP is able to act as a proteolytic activator of proMMP-2 [9]. Ten proMMPs possess a furin-like proprotein convertase recognition sequence and they are likely to be activated intracellularly and secreted or cell surface bound as active enzymes. Another unique feature of MMPs is that many of them can also be activated by non-proteolytic agents such as mercurial compounds, denaturants (both *in vitro* activation) and reactive oxygen and nitrogen species. The latter activation process may take place under inflammatory conditions [10]. Activation is thought to be induced by an oxidative modification of the thiol residue, resulting in disruption of the Cys-Zinc interaction and subsequent cleavage of the propeptide domain by autoactivation [11].

After activation, MMP activity is mainly controlled by the generic protease inhibitor α_2 -macroglobulin and TIMPs [12]. Human α_2 -macroglobulin is a plasma glycoprotein consisting of four identical subunits of 180 kDa. It inhibits most proteases by irreversible entrapment of the protease within the macroglobulin, which causes a conformational change, after which the complex is rapidly cleared by receptor (low density lipoprotein receptor-related protein-1) mediated endocytosis [13]. TIMPs are specific inhibitors of MMP activity in the tissue, and four homologues (TIMP-1 to -4) have been identified thus far [14]. TIMPs form noncovalent binding complexes with MMPs in a 1:1 stoichiometry with pro and activated forms of MMPs. The crystal structure of the complex between TIMP-1 and the catalytic domain of MMP-3 revealed that the catalytic zinc atom is bidentately chelated by Cys-1 (N-terminal amino group and carbonyl group) of TIMP-1, which expels the water molecule bound to the zinc atom [15]. Other parts (Cys1-Val4) of the TIMP-1 N-terminus bind to the active site cleft in a substrate-like manner. TIMPs inhibit all MMPs tested so far, except for MT1-MMP, MT3-MMP, MT5-MMP and MMP-19 being poorly inhibited by TIMP-1. Besides MMPs, TIMP-3 has been shown to inhibit other metzincins like ADAM-10, -12, -17 (a disintegrin and metalloprotease domain) and ADAMTS-1, -4 and -5 (a disintegrin and metalloprotease domain and thrombospondin motif). The multi-functionality of TIMPs is illustrated by many other biological activities, independent of their MMP-inhibitory activities [4].

1.1.2 Biological and pathological roles of MMPs

The proteolytic activity of MMPs is directed against most constituents of the extra cellular matrix (ECM), like collagen, elastin, fibronectin, laminin and proteoglycans and is normally tightly controlled through various regulatory events. Although various types of other proteases are also implicated in ECM degradation, MMPs are considered to be the major enzymes. The timely breakdown of ECM is essential for many normal biological processes like embryonic development, wound healing, cell migration, angiogenesis, tissue repair and remodelling and nerve growth. It is becoming increasingly clear, however, that MMPs are also involved in the release of bioactive fragments and growth factors from both ECM and also non-ECM molecules (such as other proteases, protease inhibitors, clotting factors, chemotactic molecules, latent growth factors, growth factor binding proteins, cell surface receptors, as well as cell-cell and cell-matrix adhesion molecules), thereby influencing cellular behaviour such as cell migration, differentiation, growth, inflammatory processes, apoptosis, etc [16,17]. As major sources of MMPs, immune cells rely heavily on these enzymes to mediate extravasation into tissues during inflammation. The net effect of a mixture of active MMPs is determined by their substrate specificity, cellular sources, area of distribution, and the level of regulatory proteins [18].

Under normal physiological conditions, MMP transcripts are expressed at low levels, but these levels rise rapidly when tissues are locally induced to undergo remodelling. A lack of temporal and spatial control of MMP activity can lead to excessive degradation or remodelling of ECM, which is a major concern in several pathological conditions. Critical roles for MMPs have been described in a multitude of diseases [18], including COPD (chronic obstructive pulmonary disease), asthma [19], arthritis [20], cancer [21,22], and cardiovascular disease (such as atherosclerosis and heart failure) [23,24].

1.1.3 MMP inhibitors for therapeutic intervention

The detection of different MMPs in many cancers and their association with poor prognosis has emphasized their role as potential drug targets [25,26]. In addition, studies with animal models and *in vitro* assays have indicated that broad spectrum synthetic MMP inhibitors could be effective against a variety of inflammatory diseases such as multiple sclerosis, glomerulonephritis, emphysema, aortic aneurysm and others [27]. Some synthetic hydroxamate type MMP inhibitors have been evaluated in clinical trials, mainly as treatment options for

cancer [28, 29] and rheumatoid arthritis [30]. The results of these clinical trials have mainly been disappointing, primarily because of lack of efficacy and survival benefits, and side effects such as musculoskeletal pain [28, 31]. One of the possible reasons for the failure of MMP inhibitors in clinical studies is their use in advanced stage cancers, while there still is a lack of knowledge about the role of specific MMPs in specific stages of tumor progression. Due to the absence of analytical methods to assess MMP inhibitor efficacy (by evaluating the effect on MMP activity), it remains unclear whether any compound has reached levels sufficient to inhibit target MMP activity within the tumor tissue [28]. Attempts to analyse plasma or serum levels of the gelatinases by zymography have been uninformative [32]. The use of MMP inhibitors with lower affinities for the related ADAMs showed less side effects, indicating the importance of drug selectivity [33].

1.1.4 MMP analysis

Zymography and immunological methods are mostly used to study the role of MMPs in various diseases. Although these methods can reach high sensitivities, obtaining information about the functional state of MMPs (pro-, active, or inhibited) is difficult. If the role of MMPs in disease mechanisms are to be analyzed, profiling the levels of MMP activities rather than overall MMP abundance will be necessary. This is because in pathological situations, tissue damage or other pathological effects are likely caused by active enzymes. Another complicating factor in the analysis of MMPs is the fact that samples are often dominated by pro- and inhibited-MMPs, while the most important active MMPs represent only a minor fraction. Zymography is performed under denaturing conditions and therefore does not account for key protein-protein interactions or protein-small molecule interactions which may regulate MMP activity. An immunological method based on immunocapture using MMP specific antibodies, followed by activation of a modified urokinase (containing an amino acid sequence in the pro-domain which is cleaved by MMPs), does allow activity-dependent detection of MMPs in complex samples [34]. The specific nature of this method does, however, not allow analysis on a family-wide scale, a feature which would be favourable in MMP analysis, since MMPs are often able to act in concert [35–37]. An activity-based, parallel MMP measurement would also be attractive to avoid excluding MMP members which are not known to be involved in a given disease process. Simultaneous analysis of functionally related metallo-proteases from other enzyme families would be another attractive feature.

In the field of proteomics, measurements on a proteome wide scale have become possible but most global approaches suffer from the fact that a significant part of the proteome remains undetected [38, 39]. Moreover, global approaches are only capable of measuring protein abundance, which, for the highly regulated MMP family is not correlated to activity.

The development of so-called chemical proteomics approaches relies on the use of activity-based probes, usually consisting of reactive groups for covalent interactions with the active site of the enzymes and tags for selective visualization or enrichment [40]. For MMPs and other metallo-proteases this approach was recently described using a compound that combines a hydroxamate-based inhibitor with a photo-crosslinking group and an enrichment or visualization tag [41]. With these probes, it was possible to selectively label active MMPs, added to complex proteomes. Interestingly, three non-MMP metallo-proteases were identified in invasive human melanoma cell lines, which shared no sequence homology with MMPs. This demonstrates that this inhibitor based method can be used to obtain information about potential off-targets of MMP inhibitors, which may even be related to the side-effects and disappointing results observed in the clinical trials with MMP inhibitors. Despite the potential of this approach, no endogenous active MMPs have been identified yet.

1.2 Affinity sorbents in targeted proteomics

Proteomics deals with the identification and quantification of the total protein content (the proteome) in a given sample. This comprehensive, hypothesis-generating approach, can result in the discovery of previously unknown actors in biological or pathological events. The enormous complexity of each proteome, caused by the large numbers of proteins, their concentration differences, their different physico-chemical properties and possible post-translational modifications only allows a portion of the proteome to be analyzed with 2-dimensional gel electrophoresis (2DE) [39]. The "shotgun approach", in which a digested sample is separated by multidimensional liquid chromatography prior to mass spectrometry (MS) detection, is less biased with respect to physico-chemical properties, but also has problems with low abundance proteins, because proteome complexity, which is multiplied by the digestion, exceeds the resolving power and capacity of this technique [42]. Thus sample complexity outweighs the technologies currently available for true comprehensive proteomics, preventing the detection of important low-abundance proteins like growth factors. With the use of proteome prefractionation techniques, the holistic view is sacrificed in order to access the

low-abundance proteins. Proteome prefractionation techniques include various chromatographic and electrophoretic methods where prefractionation is based on generic physico-chemical properties of proteins [43]. Prefractionation can also be based on functional protein properties, like in immunodepletion of high-abundance proteins [44], sub cellular fractionation [45], and the use of affinity methods. Affinity purification of tagged proteins and their associated proteins is another important application of affinity methods in proteomics [46].

1.2.1 Properties of affinity sorbents

Targeted proteomics is a hypothesis-driven approach where proteome prefractionation is based on selective molecular interactions between an affinity ligand or probe and proteins based on common structural or biochemical properties. The selective nature of these interactions allows high enrichment factors, resulting in detection of low-abundance proteins while additional information, for example about enzymatic activity, may also be gained. The activity-based probe (ABP) method utilizes specific chemical probes that are designed to label enzyme families in solution in an activity-dependent fashion. The probe contains a recognition element to confer specificity for the desired enzyme family, a reactive group for covalent coupling to the active site, and a tag for enrichment or detection of the labelled proteins. ABPs have been used for proteomic analysis of cysteine-, serine- and metallo-proteases and protein tyrosine phosphatases [41,47–49]. One possible drawback of the ABP approach is the bulky nature of the tag, affecting protein binding affinity, probe uptake and cellular and tissue distribution, thus hindering meaningful *in vivo* profiling experiments. This has driven the development of the so-called "click-chemistry" approach, in which proteins are first covalently labelled with an easily activatable small group, which is subsequently (e.g. within the cell), conjugated with a tag [50]. Although the ABP approach works well for small sample volumes, it is less suited for diluted large-volume samples like body fluids, which would require large amounts of expensive probes that cannot be reused.

Affinity sorbents with immobilized affinity ligands have traditionally been used as a selective purification tool in downstream processing of, for example, therapeutic proteins. In targeted proteomics, affinity sorbents are used for the enrichment and analysis of functionally related proteins. Because of its chromatographic nature, this approach is well suited for large sample volumes. Targeted proteins can be strongly concentrated while sample complexity is reduced by the removal of all

proteins that do not interact with the immobilized ligand. Additional advantages of affinity sorbents are their potential reusability and their integration with downstream analysis steps when used in column or cartridge format.

An affinity sorbent consists of a chromatographic support with immobilized affinity ligands for capturing target proteins. A support matrix with a large effective surface area and free diffusion in and out of the pores is ideal for optimal contact between affinity ligands and the targeted proteins in the sample [51]. Despite its limited pressure stability, more than two-thirds of all targeted proteomics investigations employing affinity sorbents rely on the use of agarose based materials (including cross-linked agarose) due to the wide commercial availability of preactivated supports. Reactive groups such as N-hydroxysuccinimideesters, epoxides, cyanogenbromide-, carbonyldiimidazole- or epibromohydrin-activated OH-groups are usually tethered to the support via a spacer and allow simple one- or two-step immobilization procedures under mild conditions. Other, more rarely used support materials include modified silicas [52] and vinyl polymers [53]. Although not tested with biological samples, potential advances in the support material were demonstrated with the use of polymethacrylate-based monolithic columns containing immobilized lectins [54]. Selective and strong retention of purified glycoproteins resulted in enrichment factors around 400.

Ideally the support matrix should display no non-specific protein adsorption, although this is never the case in reality. With the introduction of novel hydrophilic spacers between vinyl-based supports and the ligand, Shiyama *et al.* were able to reduce non-specific protein binding of tissue lysates [53]. Compared with an agarose-based support, however, where non-specific binding was nearly absent, the level of non-specific protein binding was still relatively high. The hydrophilic nature and absence of surface charges are responsible for the low non-specific protein binding of agarose-based supports, which is another reason for their widespread use in targeted proteomics.

The physico-chemical properties of the affinity ligand can also be a source of non-specific adsorption, which may complicate the final analysis. In an attempt to identify the primary target of an anticancer agent, Oda *et al.* prepared two affinity sorbents: one containing the active compound and the other containing a structurally closely related but inactive analogue [55]. To determine the support-mediated non-specific binding, an ethanolamine-conjugated, negative control matrix was also tested. Despite extensive washing after the enrichment with a buffer containing 0.05% surfactant (3-(3-cholamidopropyl)-dimethylammonio propane sulfonate: CHAPS), it was shown that most of the 285 enriched and identified proteins were captured in a non-specific manner, of which many were high-abundant

proteins like tubulins and molecular chaperones. A differential labeling and tagging approach was used to discriminate between specific and non-specific binding and allowed the identification of a potential target protein. Inclusion of surfactants and maybe even low levels of organic solvents to suppress non-specific adsorption during loading and washing steps can be considered, but care should be taken not to influence the structure of the targeted proteins.

Maintaining the functionality of the affinity ligand upon immobilization is another critical factor. To avoid decreased functionality upon immobilization, caused by steric hindrance, a spacer is needed. Understanding the structure-activity-relationship is crucial for successful immobilization (especially for small molecule ligands). It must be known which ligand sites are not engaged in the interaction with the target protein(s) and can serve as an attachment point for immobilization. If this site, which should be accessible, does not contain the proper functional group for immobilization, it is necessary to select or synthesize a closely related analogue with comparable binding properties and the required functional group (e.g. a primary amine). This approach has been used successfully in proteome-wide drug selectivity studies with kinase inhibitors [56].

The optimal ligand density is traditionally determined by measuring the sorbent's binding capacity for a given target protein [51]. Franco Fraguas *et al.* showed that the capacity of a lectin affinity sorbent increased with higher ligand density but that enrichment was more effective at lower ligand densities, because relatively fewer ligand molecules were required to capture a single target glycoprotein [57]. This may be an indication of increasing steric hindrance with increasing ligand density. For the most effective capturing in a targeted proteomics experiment, the total binding capacity is not a reliable parameter to determine the optimal ligand density, since affinity supports are generally not used under saturating binding conditions.

In targeted proteomics, affinity sorbents have been used for selective enrichment of various protein classes such as glutathione transferases [58], gelatin-binding proteins [59] and heparin-binding proteins [60]. The majority of research has focused, however, on kinases and glycoproteins and an overview of these examples will be given here.

1.2.2 Kinase inhibitor affinity sorbents

In traditional drug development, inhibitor selectivity is only tested with limited panels of proteins, resulting in incomplete selectivity profiles. Affinity sorbents

with immobilized inhibitors can be used for the mapping of drug-protein interactions on a proteome-wide scale. With this drug target proteomics approach, useful information about potential "off target" interactions or potential alternative drug targets for new therapeutic approaches has been obtained for the class of kinases (800 estimated members). The majority of kinase drugs interact with the relatively conserved ATP binding pocket, making the synthesis of selective kinase inhibitors very difficult, and proper assessment of selectivity very important. An overview of several drug target investigations with kinase inhibitor affinity sorbents is given in Table 1.1.

Among the different targeted proteomics studies there is quite some variation in the experimental set-up: sometimes samples are passed through a column [61, 62, 66] but most studies use batch enrichment, where the affinity sorbent is incubated up to several hours with the (usually cooled) sample. Elution can be performed with competing ligands or with SDS sample buffer (sometimes with boiling), or a combination of both [56, 65, 67], to discriminate between weak- and tight-binding proteins. The downstream analytical steps vary but separation by gel electrophoresis (both 1DE and 2DE), proteolytic digestion and MS analysis are the key techniques.

An interesting approach was used by Ding *et al.*, who first screened kinase inhibitor libraries for induction of neurogenesis in stem cells and then identified GSK-3 β as the primary target of the selected inhibitor through drug target proteomic screening [68]. Inhibitor affinity sorbents can thus also be used to get more insight into molecular mechanisms.

Several papers demonstrated that despite a good *in vitro* selectivity (tested with kinase panels), unexpected other kinases and protein targets could be identified. This shed new light on possible side effects of the drugs or on drug resistance and may open opportunities for further optimization of the compounds. The identification of targets with a dramatic affinity difference from sub-nanomolar up to micromolar makes it difficult, however, to obtain quantitative information about inhibition potency. It is thus necessary to perform *in vitro* assays with purified targets to confirm the initial findings.

The varying numbers of identified novel targets, shown in Table 1.1, is likely caused by the use of different inhibitor affinity sorbents, the screening of different proteomes and differences in the analytical steps. The identification of just one or a few targets in some papers gives the impression of high drug selectivity [64, 68]. Considering the relatively conserved ATP binding pocket and the large number of kinases, together with the identification of several targets with a broad range of affinities in other papers [63], this apparently high selectivity can be questioned.

Immobilized drug	Screened proteomes	Known targets	Novel targets	Findings	Ref.
Flavopiridol	HeLa, Prostate, A549 and NSCLC cells	Cyclin-dependent kinases (CDKs)	Cytosolic aldehyde dehydrogenase (ALDH-1), glycogen phosphorylase (GP)	Suggested ALDH-1 and GP-dependent drug resistance	[61] [62]
Purvalanol B and inactive derivative	Diverse parasite cell types	CDKs	Several unexpected kinases	Selectivity lower than expected Low affinity kinases enriched	[63]
Gwenpaullone	Tissue extracts and parasites	CDKs and glycogen synthase kinase-3 (GSK-3)	Mitochondrial malate dehydrogenase (MDH)	Major targets: GSK-3 α/β , MDH dependent parasite toxicity	[63]
Hymenialdisine and analogues	Mouse brain	CDKs, GSK-3 β , CK1	p90RSK	Only 1 novel target	[64]
PHA-539136 (cyclopropyl-pyrazole class)	HCT 116 and Pancreatic acinar cells	CDK2	14 non-kinases	Two heat shock proteins have strong <i>in vitro</i> inhibitor affinity	[65]
SB203580 analogue (pyridinyl imidazole class)	HeLa and COS-7 cells	p38	Several kinases and other proteins	Low drug selectivity Novel kinase targets confirmed with <i>in vitro</i> assays	[66]
Bisindolylmaleimide analogues	HeLa, COS-7, HuH-7 cells	Protein kinase C isozymes GSK-3 α/β , Rsk2	Ste20-related kinase, CDK2, adenosine kinase, quinon reductase type 2	Low affinity kinases enriched	[67]
PP58 (pyrido[2,2-d]pyrimidine class)	HeLa cells	Protein tyrosine kinase	>30 kinases	Non tyrosin-kinases potentially inhibited <i>in vitro</i> Low affinity kinases enriched	[56]
TWS119 (4,6-disubstituted pyrrolopyrimidine)	Mouse P19 EC stem cells	No	GSK-3 β	Only 1 novel target GSK-3 β involved in stem cell neurogenesis	[68]

Table 1.1: Examples of drug target proteomic studies with kinase inhibitor affinity sorbents.

Drug target screening with hymenialdisine (a sponge-derived, natural product kinase inhibitor) as ligand resulted in identification of only one novel target next to three known ones, which is indeed less than may be expected, considering that IC_{50} values below 10 $\mu\text{mol/L}$ were measured for 39 out of a panel of 60 tested kinases [64]. Another possible explanation for this finding is a relatively high abundance of the identified kinases while other kinases were present at too low levels for detection, even after enrichment. Despite some very interesting results, drug-target proteomics technology may still miss relevant proteins of very low abundance but with high affinities due to the limited sensitivity of the analytical methodologies used, especially when gel electrophoresis is involved.

1.2.3 Glycoproteomics

Glycosylation is one of the most common but also one of the most complex among the various types of post-translational modifications. It modulates protein function and plays fundamental roles in diverse processes such as the immune response and cellular recognition. Glycosylation patterns on the same protein can vary with cell type, growth rate of cells (e.g. due to transformation), development, type of disease and even with disease progression. Alterations in glycosylation of cell surface proteins are, for example, a hallmark of cancer [69]. Lectin affinity sorbents are effective materials to enrich glycoproteins and glycopeptides based on defined carbohydrate structures on a proteome-wide scale. Binding specificity is achieved by a highly selective interaction of the carbohydrate moiety with the binding pocket of the lectin. Glycoproteomic methods using lectin affinity sorbents can be divided into methods employing tryptic digestion before or after enrichment. Most methods rely on mass spectrometry for identification of the enriched glycoproteins or glycopeptides. A common feature of nearly all glycoproteomic studies is the use of deglycosylation enzymes prior to mass spectrometry to remove the oligosaccharides and to allow identification of the core peptide structures by comparison with sequences in DNA and protein databases.

Digestion before enrichment simplifies the downstream analysis workflow, because all non-glycosylated peptides are removed during lectin enrichment. Regnier *et al.* developed two approaches where affinity enrichment is combined with stable isotopic labeling for comparative quantification [52, 70, 71]. In one approach, two digested samples are differentially labeled with stable isotope coding reagents, followed by mixing and affinity enrichment with a Lotus tetragonolobus agglutinin (LTA) affinity sorbent, which is specific for glycopeptides carrying α -L-Fucose. After deglycosylation, the peptides were fractionated (RP-HPLC) and analyzed

by MALDI-MS. This method revealed that a series of fucosylated serum proteins decreased more than 2-fold during chemotherapy. These proteins were known to be involved in cell adhesion and cancer cell migration [52].

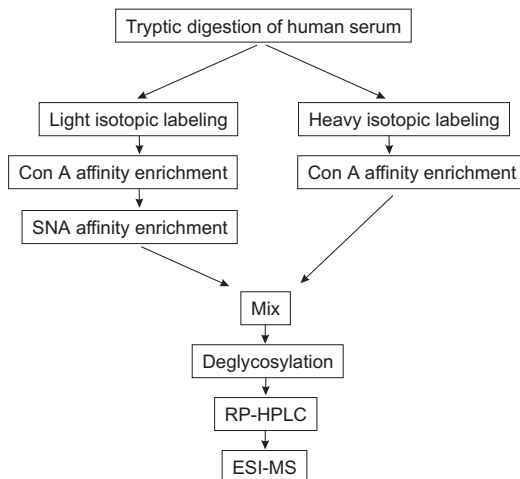


Figure 1.1: Strategy to quantify the extend of sialylation of N-linked glycopeptides by combining the broad selectivity of the Con A with the narrow selectivity of the SNA lectin sorbent. Aliquots of a digested serum sample are labelled with either a light or a heavy isotopic tag and glycopeptides are enriched with the Con A lectin sorbent. In the next step only the sialylated peptides (labeled with light tag) are enriched with the SNA lectin sorbent and mixed with the glycopeptides (containing both the sialylated and non-sialylated glycopeptides), labelled with the heavy tag. Both aliquots are mixed and deglycosylated to remove glycan heterogeneity. After fractionation with RP-HPLC, the degree of sialylation is determined by measuring the isotopic ratio with ESI-MS.

The extend of protein sialylation was assessed by an advanced method comprising differential isotopic labeling combined with both serial and parallel lectin affinity enrichment (using both Con A and SNA affinity sorbent; see Fig. 1.1). The majority of studied peptides appeared to be fully sialylated [70].

Yang and Hancock developed a comprehensive glycoproteomics method by using a mixed immobilized lectin column containing 3 different immobilized lectins, with complementary specificities for carbohydrate motifs, covering the most common O- and N-linked glycans in serum proteins [72, 73]. Captured glycoproteins were sequentially eluted with different sugars competing with the lectin carbohydrate binding site. Changes in the distribution of glycoproteins in the elution fractions were proposed to be an indication for a change in glycosylation pattern.

Membrane proteins, which are notoriously difficult to analyze by proteomics techniques, and proteins which are shed from the membrane are often extensively glycosylated and can therefore be investigated with lectin-based glycoproteomic methods. Guo *et al.* investigated shedding by culturing a cell line with and without a metalloprotease inhibitor [74]. The value of the lectin enrichment step with a wheat germ agglutinin (WGA) sorbent was clearly demonstrated by the fact that differences between the cell supernatant with or without the inhibitor could be revealed by 2DE after enrichment only. Differential isotopic labelling was used for relative quantification allowing identification of three proteins that were shed in a metalloprotease-dependent manner. TACE (tumor necrosis factor- α converting enzyme) was shown to be the responsible protease in experiments with TACE +/+ and -/- cells.

Another study with ConA and WGA lectin affinity sorbents demonstrated a 150-200 fold enrichment of membrane glycoproteins [75]. The majority of the 158 identified proteins was uniquely enriched with one of the lectin affinity sorbents. The physico-chemical properties of the identified proteins were very diverse, indicating that lectin enrichment is not biased towards parameters such as size and charge.

A different type of glycoproteomics was performed with an immobilized mannose affinity sorbent to study mannose binding proteins in the rice proteome [76]. From different rice tissues, 136 distinct mannose binding proteins were identified, of which only 36% were previously known to be involved in sugar metabolism.

1.2.4 Inhibitor affinity sorbents for activity-based MMP analysis

The affinity sorbent approach in targeted proteomics has emerged as an important tool to reduce sample complexity and to access low-abundance proteins, which are difficult to detect using comprehensive proteomic approaches. Several papers report that considerable enrichment factors can be reached due to the high selectivity of the affinity ligands and the option to concentrate samples strongly. Compared to the ABP approach, which is more convenient to perform by just adding the label to the sample, affinity sorbents can also be used for proteins without enzymatic activity and their potential use is thus broader. Another advantage, which has not been exploited yet, is the automation potential that affinity sorbents offer when used in robotic solid-phase extraction platforms. Automation is likely needed to reach the necessary reproducibility for studies with

larger sample sets.

The aim of this thesis is to develop an integrated system for activity-based profiling of MMPs, using an immobilized inhibitor affinity sorbent for activity-based MMP enrichment, which is coupled on-line to digestion, separation and mass spectrometric detection steps.

In the enrichment step of this method, immobilized synthetic reversible MMP inhibitors are chosen as affinity ligands, for several reasons. First, synthetic inhibitors are stable and relatively easy to immobilize using inhibitor functionalities which are not involved in the interaction with the enzymes to be captured from the sample. Second, the selective nature of the inhibitor affinity chromatography approach (compared to other modes of chromatography like reversed phase or ion exchange) is a requirement to obtain high enrichment factors, needed to reach the desired sensitivity to detect low-abundance MMPs. Third, depending on the relative affinity for the target and contact time of extraction, the immobilized, active site binding inhibitors should selectively enrich active MMPs (something which is difficult to achieve with, for example, immobilized antibodies). Because of active site inaccessibility, inactive proMMPs and inhibitor-MMP complexes should not be captured, which would result in an activity-based enrichment. Fourth, broad-spectrum inhibitors may be able to capture multiple MMPs, omitting the need to know in advance which MMPs can be expected (as is the case with most immunological methods) and increasing the chance to discover new participants in pathological or physiological events. Immobilized hydroxamid acid-based MMP inhibitors have already been shown to be effective ligands in affinity column purifications of MMP-1, -7 and -8 [77–79]. And fifth, the selectivity of the immobilized inhibitor can be assessed on a proteome-wide scale, which may result in the identification of new targets and off-targets, giving insight in new therapeutic strategies and side effects (observed in clinical trials).

The inhibitors are immobilized on a solid support because a highly effective affinity ligand concentration can be reached this way, favouring enzyme capture. This may be needed to capture low abundance-active MMPs from large volume samples (like body fluids). Another advantage of the solid support approach is the potential for automation, after packing the inhibitor affinity sorbent in a cartridge, and the option for on-line coupling to downstream processing and analytical steps.

The enrichment step is coupled on-line to digestion and analytical steps to avoid the use of labour intensive and time consuming gel electrophoresis steps. On-line digestion with immobilized trypsin was chosen because of its favourable kinetics compared to in-solution or in-gel digestion. For final analysis of the tryptic MMP peptides, the two automated sample preparation steps were coupled on-line to

liquid chromatography mass spectrometry (LC-MS) for separation and identification of tryptic MMP peptides.

From a broader scientific perspective, it is interesting to investigate if integrated protein analysis systems can contribute to one of the remaining difficulties in proteomics, which is the analysis of low-abundance proteins. Although integrated systems have been described in the literature, the challenging goal of analyzing low-abundance proteins with this approach has rarely been tackled.

1.3 Outline of the thesis

The thesis focuses on different aspects of the development of an integrated system for activity-based profiling of matrix metallo-proteases. Throughout the thesis, purified recombinant MMP-12 (catalytic domain) is used as a model enzyme. The feasibility of activity-based MMP-12 enrichment with an immobilized inhibitor affinity sorbent, using batch extractions is described in **Chapter 2**. An inhibitor affinity sorbent with an immobilised broad-range inhibitor with micromolar MMP affinity is characterised by determining the functionality of the ligand after immobilization and determining the kinetics of batch MMP-12 extractions. The extraction yields of MMP-12 added to extraction buffer are quantified indirectly using an activity assay. The activity-based nature of MMP-12 enrichment is investigated with titrations of the two endogenous MMP-12 inhibitors, namely TIMP-1 and α_2 -macroglobulin. The enrichment of MMP-12 spiked at a low level into a complex sample matrix is also investigated. **Chapter 3** considers the automation of the extraction which is needed to couple it to downstream analytical steps. Automated extractions are performed using a solid-phase extraction (SPE) platform and inhibitor affinity sorbent packed into cartridges. Extraction yields are investigated at different flow rates and with different MMPs. A nanomolar affinity inhibitor is also tested as affinity ligand and compared with the micromolar inhibitor affinity sorbent. **Chapter 4** describes the chemical modification of immobilized trypsin and its positive effects on trypsin autolysis and digestion kinetics of the model protein cytochrome c. **Chapter 5** deals with the implementation of the extraction cartridge developed in Chapter 3 and the digestion cartridge developed in Chapter 4 into an integrated system for activity-based profiling of MMPs. The on-line coupling of the two sample preparation steps to LC-MS is described together with the effect of surfactant addition to extraction and digestion buffers. Finally, the system performance is evaluated with MMP-12 spiked into urine samples.

References

- [1] W. Stocker, F. Grams, U. Baumann, P. Reinemer, F. X. Gomis-Ruth, D. B. McKay, W. Bode; *Protein Sci* **4**, 823 (1995)
- [2] J. Gross, C. M. Lapière; *Proc Natl Acad Sci U S A* **48**, 1014 (1962)
- [3] H. Nagase, R. Visse, G. Murphy; *Cardiovasc Res* **69**, 562 (2006)
- [4] H. Nagase, J. Woessner, J. F.; *J Biol Chem* **274**, 21491 (1999)
- [5] J. A. Kwan, C. J. Schulze, W. Wang, H. Leon, M. Sariahmetoglu, M. Sung, J. Sawicka, D. E. Sims, G. Sawicki, R. Schulz; *Faseb J* **18**, 690 (2004)
- [6] G. A. Limb, K. Matter, G. Murphy, A. D. Cambrey, P. N. Bishop, G. E. Morris, P. T. Khaw; *Am J Pathol* **166**, 1555 (2005)
- [7] D. Luo, B. Mari, I. Stoll, P. Anglard; *J Biol Chem* **277**, 25527 (2002)
- [8] H. E. Van Wart, H. Birkedal-Hansen; *Proc Natl Acad Sci U S A* **87**, 5578 (1990)
- [9] Y. Itoh, M. Seiki; *J Cell Physiol* **206**, 1 (2006)
- [10] Z. Gu, M. Kaul, B. Yan, S. J. Kridel, J. Cui, A. Strongin, J. W. Smith, R. C. Liddington, S. A. Lipton; *Science* **297**, 1186 (2002)
- [11] K. K. Nelson, J. A. Melendez; *Free Radic Biol Med* **37**, 768 (2004)

- [12] K. Brew, D. Dinakarpanadian, H. Nagase; *Biochim Biophys Acta* **1477**, 267 (2000)
- [13] S. J. Kolodziej, T. Wagenknecht, D. K. Strickland, J. K. Stoops; *J Biol Chem* **277**, 28031 (2002)
- [14] D. E. Gomez, D. F. Alonso, H. Yoshiji, U. P. Thorgeirsson; *Eur J Cell Biol* **74**, 111 (1997)
- [15] F. X. Gomis-Ruth, K. Maskos, M. Betz, A. Bergner, R. Huber, K. Suzuki, N. Yoshida, H. Nagase, K. Brew, G. P. Bourenkov, H. Bartunik, W. Bode; *Nature* **389**, 77 (1997)
- [16] J. D. Mott, Z. Werb; *Curr Opin Cell Biol* **16**, 558 (2004)
- [17] I. Stamenkovic; *J Pathol* **200**, 448 (2003)
- [18] K. A. Szabo, R. J. Ablin, G. Singh; *Clin Applied Immunol Rev* **4**, 295 (2004)
- [19] I. K. Demedts, G. G. Brusselle, K. R. Bracke, K. Y. Vermaelen, R. A. Pauwels; *Curr Opin Pharmacol* **5**, 257 (2005)
- [20] I. Tchetverikov, H. K. Ronday, B. Van El, G. H. Kiers, N. Verzijl, J. M. TeKoppele, T. W. Huizinga, J. DeGroot, R. Hanemaaijer; *Ann Rheum Dis* **63**, 881 (2004)
- [21] R. Ala-aho, V. M. Kahari; *Biochimie* **87**, 273 (2005)
- [22] M. Polette, B. Nawrocki-Raby, C. Gilles, C. Clavel, P. Birembaut; *Crit Rev Oncol Hematol* **49**, 179 (2004)
- [23] A. C. Newby; *Cardiovasc Res* **69**, 614 (2006)
- [24] F. G. Spinale; *Circ Res* **90**, 520 (2002)
- [25] A. Vaisanen, M. Kallioinen, P. J. Taskinen, T. Turpeenniemi-Hujanen; *J Pathol* **186**, 51 (1998)
- [26] A. J. Minn, G. P. Gupta, P. M. Siegel, P. D. Bos, W. Shu, D. D. Giri, A. Viale, A. B. Olshen, W. L. Gerald, J. Massague; *Nature* **436**, 518 (2005)
- [27] C. E. Brinckerhoff, L. M. Matrisian; *Nat Rev Mol Cell Biol* **3**, 207 (2002)
- [28] L. M. Coussens, B. Fingleton, L. M. Matrisian; *Science* **295**, 2387 (2002)
- [29] S. Zucker, J. Cao, W. T. Chen; *Oncogene* **19**, 6642 (2000)

- [30] K. S. Mix, M. B. Sporn, C. E. Brinckerhoff, D. Eyre, D. J. Schurman; *Clin Orthop Relat Res* **S129–S137** (2004)
- [31] A. H. Baker, D. R. Edwards, G. Murphy; *J Cell Sci* **115**, 3719 (2002)
- [32] M. Hidalgo, S. G. Eckhardt; *J Natl Cancer Inst* **93**, 178 (2001)
- [33] P. D. Brown; *Expert Opin Investig Drugs* **9**, 2167 (2000)
- [34] R. Hanemaaijer, H. Visser, Y. T. Konttinen, P. Koolwijk, J. H. Verheijen; *Matrix Biol* **17**, 657 (1998)
- [35] U. B. Hofmann, R. Houben, E. B. Brocker, J. C. Becker; *Biochimie* **87**, 307 (2005)
- [36] M. A. Moses, D. Wiederschain, K. R. Loughlin, D. Zurakowski, C. C. Lamb, M. R. Freeman; *Cancer Res* **58**, 1395 (1998)
- [37] P. Thomas, R. Khokha, F. A. Shepherd, R. Feld, M. S. Tsao; *J Pathol* **190**, 150 (2000)
- [38] G. L. Corthals, V. C. Wasinger, D. F. Hochstrasser, J. C. Sanchez; *Electrophoresis* **21**, 1104 (2000)
- [39] S. P. Gygi, G. L. Corthals, Y. Zhang, Y. Rochon, R. Aebersold; *Proc Natl Acad Sci U S A* **97**, 9390 (2000)
- [40] D. A. Campbell, A. K. Szardenings; *Curr Opin Chem Biol* **7**, 296 (2003)
- [41] A. Saghatelian, N. Jessani, A. Joseph, M. Humphrey, B. F. Cravatt; *Proc Natl Acad Sci U S A* **101**, 10000 (2004)
- [42] M. P. Washburn, D. Wolters, I. Yates, J. R.; *Nat Biotechnol* **19**, 242 (2001)
- [43] P. G. Righetti, A. Castagna, P. Antonioli, E. Boschetti; *Electrophoresis* **26**, 297 (2005)
- [44] R. Pieper, Q. Su, C. L. Gatlin, S. T. Huang, N. L. Anderson, S. Steiner; *Proteomics* **3**, 422 (2003)
- [45] S. Brunet, P. Thibault, E. Gagnon, P. Kearney, J. J. Bergeron, M. Desjardins; *Trends Cell Biol* **13**, 629 (2003)

- [46] A. C. Gavin, M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J. M. Rick, A. M. Michon, C. M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M. A. Heurtier, R. R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, G. Superti-Furga; *Nature* **415**, 141 (2002)
- [47] B. F. Cravatt, E. J. Sorensen; *Curr Opin Chem Biol* **4**, 663 (2000)
- [48] D. A. Jeffery, M. Bogyo; *Curr Opin Biotechnol* **14**, 87 (2003)
- [49] S. Kumar, B. Zhou, F. Liang, W. Q. Wang, Z. Huang, Z. Y. Zhang; *Proc Natl Acad Sci U S A* **101**, 7943 (2004)
- [50] A. E. Speers, B. F. Cravatt; *Chem Biol* **11**, 535 (2004)
- [51] S. R. Narayanan, L. J. Crane; *Trends Biotechnol* **8**, 12 (1990)
- [52] L. Xiong, D. Andrews, F. Regnier; *J Proteome Res* **2**, 618 (2003)
- [53] T. Shiyama, M. Furuya, A. Yamazaki, T. Terada, A. Tanaka; *Bioorg Med Chem* **12**, 2831 (2004)
- [54] M. Bedair, Z. El Rassi; *J Chromatogr A* **1079**, 236 (2005)
- [55] P. H. O'Farrell; *J Biol Chem* **250**, 4007 (1975)
- [56] J. Wissing, K. Godl, D. Brehmer, S. Blencke, M. Weber, P. Habenberger, M. Stein-Gerlach, A. Missio, M. Cotten, S. Muller, H. Daub; *Mol Cell Proteomics* **3**, 1181 (2004)
- [57] L. Franco Fraguas, F. Batista-Viera, J. Carlsson; *J Chromatogr B Analyt Technol Biomed Life Sci* **803**, 237 (2004)
- [58] A. J. van Rossum, P. M. Brophy, A. Tait, J. Barrett, J. R. Jefferies; *Proteomics* **1**, 1463 (2001)
- [59] M. Saito, M. Kimoto, T. Araki, Y. Shimada, R. Fujii, K. Oofusa, M. Hide, T. Usui, K. Yoshizato; *Eur Urol* **48**, 865 (2005)
- [60] K. Shefcheck, X. Yao, C. Fenselau; *Anal Chem* **75**, 1691 (2003)
- [61] A. Kaiser, K. Nishi, F. A. Gorin, D. A. Walsh, E. M. Bradbury, J. B. Schnier; *Arch Biochem Biophys* **386**, 179 (2001)

- [62] J. B. Schnier, G. Kaur, A. Kaiser, S. F. Stinson, E. A. Sausville, J. Gardner, K. Nishi, E. M. Bradbury, A. M. Senderowicz; *FEBS Lett* **454**, 100 (1999)
- [63] M. Knockaert, K. Wieking, S. Schmitt, M. Leost, K. M. Grant, J. C. Mottram, C. Kunick, L. Meijer; *J Biol Chem* **277**, 25493 (2002)
- [64] Y. Wan, W. Hur, C. Y. Cho, Y. Liu, F. J. Adrian, O. Lozach, S. Bach, T. Mayer, D. Fabbro, L. Meijer, N. S. Gray; *Chem Biol* **11**, 247 (2004)
- [65] G. Lolli, F. Thaler, B. Valsasina, F. Roletto, S. Knapp, M. Uggeri, A. Bachi, V. Matafora, P. Storici, A. Stewart, H. M. Kalisz, A. Isacchi; *Proteomics* **3**, 1287 (2003)
- [66] K. Godl, J. Wissing, A. Kurtenbach, P. Habenberger, S. Blencke, H. Gutbrod, K. Salassidis, M. Stein-Gerlach, A. Missio, M. Cotten, H. Daub; *Proc Natl Acad Sci U S A* **100**, 15434 (2003)
- [67] D. Brehmer, K. Godl, B. Zech, J. Wissing, H. Daub; *Mol Cell Proteomics* **3**, 490 (2004)
- [68] S. Ding, T. Y. Wu, A. Brinker, E. C. Peters, W. Hur, N. S. Gray, P. G. Schultz; *Proc Natl Acad Sci U S A* **100**, 7632 (2003)
- [69] G. F. Springer; *J Mol Med* **75**, 594 (1997)
- [70] R. Qiu, F. E. Regnier; *Anal Chem* **77**, 2802 (2005)
- [71] L. Xiong, F. E. Regnier; *J Chromatogr B Analyt Technol Biomed Life Sci* **782**, 405 (2002)
- [72] Z. Yang, W. S. Hancock; *J Chromatogr A* **1053**, 79 (2004)
- [73] Z. Yang, W. S. Hancock; *J Chromatogr A* **1070**, 57 (2005)
- [74] L. Guo, J. R. Eisenman, R. M. Mahimkar, J. J. Peschon, R. J. Paxton, R. A. Black, R. S. Johnson; *Mol Cell Proteomics* **1**, 30 (2002)
- [75] D. Ghosh, O. Krokhin, M. Antonovici, W. Ens, K. G. Standing, R. C. Beavis, J. A. Wilkins; *J Proteome Res* **3**, 841 (2004)
- [76] N. L. Andon, D. Eckert, I. Yates, J. R., P. A. Haynes; *Proteomics* **3**, 1270 (2003)
- [77] W. M. Moore, C. A. Spilburg; *Biochemistry* **25**, 5189 (1986)

- [78] A. R. Welch, C. M. Holman, M. F. Browner, M. R. Gehring, C. C. Kan, H. E. Van Wart; *Arch Biochem Biophys* **324**, 59 (1995)
- [79] K. A. Mookhtiar, H. E. Van Wart; *Biochemistry* **29**, 10620 (1990)